

T.C.

DOĞUŞ ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

BİLGİSAYAR VE ENFORMASYON BİLİMLERİ

ANABİLİM DALI

**Kredi Karti Dolandiricilik Tespitinin Makine**

**Öğrenmesi Yöntemleri İle Tespiti**

Yüksek Lisans Projesi

Proje Hazırlayan

**Ahmed essam abdelghaffar ELSAYED**

**ÖĞRENCİ NO**

**2022-910-1000-3**

Danışman

**Dr. Öğr. Üyesi Zeynep Behrin GÜVEN AYDIN**

**İstanbul, 2024**



T.C.

DOĞUŞ ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

BİLGİSAYAR VE ENFORMASYON BİLİMLERİ

ANABİLİM DALI

**Kredi Karti Dolandiricilik Tespitinin Makine**

**Öğrenmesi Yöntemleri İle Tespiti**

Yüksek Lisans Projesi

Proje Hazırlayan

**Ahmed essam abdelghaffar ELSAYED**

**ÖĞRENCİ NO**

**2022-910-1000-3**

Danışman

**Dr. Öğr. Üyesi Zeynep Behrin © AYDIN**

**İstanbul, 2024**

**ÖNSÖZ**

Bu Proje internet bankacılığı dolandırıcılığıyla mücadele amacıyla tasarlanan bir sistem üzerine odaklanmaktadır, ve kullanılan makine öğrenimi algoritmalarının etkinliğini değerlendirmektedir.

**ÖZET**

Bu proje çalışması, internet bankacılığı dolandırıcılığına karşı bir sistem tasarlamayı ve değerlendirmeyi amaçlamaktadır. 2004-2007 tarihleri arasındaki müşteri işlemlerini içeren BankSim veri seti üzerinde, İleri Beslemeli Sinir Ağı, Gaussian Naive Bayes, K-Nearest Neighbors ve XGBoost gibi çeşitli makine öğrenimi algoritmaları Python programlama dili kullanılarak uygulanmıştır.. Yapılan analizler sonucunda, özellikle K-Nearest Neighbors ve XGBoost'un yüksek doğruluk ve düşük hata oranları ile dolandırıcılık tespiti konusunda etkili olduğu belirlenmiştir. Bu çalışma, internet bankacılığı dolandırıcılığıyla mücadelede kullanılabilecek etkili bir sistem önerisini sunmaktadır.

**ABSTRACT**

This project aims to design and evaluate a system to counter internet banking fraud. Utilizing the BankSim dataset spanning from 2004 to 2007, various machine learning algorithms, including Artificial Neural Network, Gaussian Naive Bayes, K-Nearest Neighbors, and XGBoost, were implemented using the Python programming language. Through the conducted analyses, it was determined that particularly K-Nearest Neighbors and XGBoost are effective in fraud detection, exhibiting high accuracy and low error rates. This study proposes an effective system for combating internet banking fraud.

**İÇİNDEKİLER**

**Sayfa No.**

**ÖNSÖZ…………………………………………………………………………………..iii**

**ÖZET ……………………………………………………………………………………iv**

**ABSTRACT……………………………………………………………………………...v**

**ŞEKİLLER LİSTESİ..................................................................................................... vii**

**TABLO LISTESI……………………………………………………………………….vii**

**KISALTMALAR………………………………………………………………………..ix**

**GİRİŞ................................................................................................................................ 1**

**LİTERATÜR ARAŞTIRMASI........................................................................................ 2**

**VERİSETİ.......................................................................................................................... 4**

**KULLANLAN TEKNİKLER........................................................................................... 5**

**ALGORITMALAR UYGULAMASI …………………………………………………. 6**

**SONUÇ………………………………………………………………………………….. 13**

**KAYNAKÇA......................................................................................................................14**

**ŞEKILLER LISTESI**

**Sayfa No.**

Şekil 1. Veri seti ilk 5 satır incelemsi ………………………………………………………6

Şekil 2. Veri seti genel bilginin incelemesi…………………………………………………7

Şekil 3. Dolandırıcılık olayların sayısı Belirtmesi………………………………………….8

Şekil 4. Dolandırıcılık olaylarının hangi kategoride daha sık gerçekleştiğini incelemesi….9

Şekil 5. Dolandırıcılık olaylarının hangi kategoride daha sık gerçekleştiğini incelemesi….9

Şekil 6. Dolandırıcılık olaylarının hangi miktarlarda daha sık gerçekleştiğini incelemesi..10

**Tablo Listesi**

**Sayfa No.**

Tablo (1) Kullanılan Algoritmalar Sonuçları………………………………………………………….. 12

**KISALTMALAR**

**ANN**: Artificial Neural Network

**GNB**: Gaussian Naive Bayes

**KNN**: K-Nearest Neighbors

**XGBoost**:eXtreme Gradient Boostin

**GİRİŞ**

Günümüzde teknolojik gelişmelerin hız kazanmasıyla birlikte, birçok alanda hayatımıza giren yenilikler insanların günlük yaşantısını önemli ölçüde etkilemektedir. Bu teknolojik ilerlemeler, özellikle kredi kartları gibi finansal araçlar aracılığıyla insanlara sağlanan kolaylıklar ve avantajlar ile birlikte, e-perakende ve e-ticaret gibi dijital platformların gelişimine de zemin hazırlamıştır. Ancak, bu gelişmelerin beraberinde getirdiği faydaların yanı sıra, kredi kartı dolandırıcılığı gibi riskleri de beraberinde taşımaktadır.

Kredi kartları, nakit taşıma zorunluluğunu ortadan kaldırarak harcamaları kolaylaştırmanın yanı sıra, internet üzerinden alışverişin yaygınlaşmasıyla birlikte e-perakende sektörünün büyümesine de katkı sağlamıştır. Ancak, bu teknolojik gelişmelerle birlikte kredi kartı dolandırıcılığı gibi güvenlik sorunları da artış göstermiştir. Bu noktada, kredi kartı kullanıcılarının ve işletmelerin maddi ve manevi kayıplar yaşamamak adına alacakları önlemler büyük önem taşımaktadır.

Öte yandan, e-perakende ve e-ticaret sektöründe yaşanan hızlı büyüme ve rekabet, işletmeleri müşteri taleplerini doğru bir şekilde karşılamak konusunda zorlayıcı hale getirmiştir. Talep tahmini, bu sektörde başarılı bir işletme yönetimi için kritik bir öneme sahiptir. Ancak, doğru talep tahmininde bulunamamanın sonuçları arasında stok fazlası, müşteri memnuniyetsizliği ve maddi kayıplar yer almaktadır.

E-ticaretin gelişimi ise internetin ve mobil teknolojilerin yaygınlaşmasıyla büyük bir ivme kazanmıştır. İşletmeler, veri bilimi ve yapay zekâ gibi teknolojik araçları kullanarak müşteri ihtiyaçlarını daha iyi anlamak ve özel ürün ve hizmetler sunmak konusunda başarılı olmaktadır. Veri odaklı kararlar, işletmelerin karlılığını artırırken, müşteri memnuniyetini de sağlamaktadır.

Bu proje kapsamında, kredi kartı dolandırıcılığına karşı güvenlik önlemleri, e-perakende sektöründe talep tahmini yöntemleri ve e-ticaret işletmelerinin veri odaklı karar alma süreçleri gibi konular incelenecektir. Bu detaylı analizler, bireylerin ve işletmelerin teknolojik gelişmelerden en iyi şekilde faydalanmalarına katkı sağlayacaktır.

**LİTERATÜR ARAŞTIRMASI**

(ÖZBAY E. , 2007) Bu tez çalışması, internet bankacılığı dolandırıcılık işlemlerini önlemek için bir sistem tasarlamayı amaçlamıştır. 2004-2007 tarihleri arasında 571 müşteriye ait 12773 parasal işlem incelenmiştir. Müşteri bilgilerine göre 16 grup oluşturulmuş, bu gruplardan biri olan grup52'de 6910 yasal ve 466 suçlu kayıt tespit edilmiştir.

Toplam 7376 veri, C4.5 ve Naive Bayes Sınıflandırıcılarına tabi tutulmuştur. Saf Bayes'e doğrudan sokulduğunda 141 doğru, 54 yanlış sınıflandırma yapılmıştır. Evet ve hayır için ayrı hesaplamalarla 195 örneğin 7'si yanlış, 188'i doğru sınıflandırılmıştır.

Sonuçlar grafiksel olarak sunulmuş, ardından sisteme girilen veriler ve ara sonuçlar detaylı bir şekilde açıklanmıştır. Sistem adım adım anlatılarak elde edilen sonuçlar, beklenen sonuçlarla karşılaştırılmış ve %0,7 oranında hata ile sonuçlandığı belirtilmiştir.

Çalışma, sistemin suçlu işlemleri belirlemek için eğitim kümesinin önemini vurgulamıştır. İlk olarak internet bankacılığına odaklanmış olsa da, küçük değişikliklerle diğer alanlarda da etkili olabileceği öne sürülmüştür. [1]  
 (ÖZMEN E. P. , ÖZCAN T. ) Çalışmada, kredi kartı dolandırıcılığının tespiti için klasik CART algoritması ve genetik algoritma ile optimize edilmiş bir GA-CART modeli değerlendirilmiştir. GA-CART modeli, %37 oranında daha yüksek doğruluk oranlarına sahiptir. Klasik CART %64,28, GA-CART ise %87,95 doğruluk oranlarına ulaşmıştır. Bu sonuçlar, GA-CART modelinin daha etkin ve doğru bir sınıflandırma sağladığını göstermektedir. Çalışma, finans sektöründe meşru olmayan davranışların tespiti için daha etkili bir araç sunmaktadır. [2]

(AYDIN T. , Dal D. , Mayıs 2021) Bu tezde, kredi kartı sahtekârlığının yapay zekâ ile tespiti için Çoklu Öznitelik Seçimi (ÇOKS) adlı yeni bir algoritmanın literatüre kazandırıldığı belirtilmektedir. Araştırma, bankacılık ve ödeme sistemlerinin güvenliğine katkı sağlamayı hedeflemektedir. ÇOKS algoritması, büyük bir veri kümesi üzerinde eğitilmiş ve elde edilen modelin testleri gerçekleştirilmiştir. Yapılan karşılaştırmalar, ÇOKS algoritmasının %99,93 doğruluk, %95,60 kesinlik ve %80,0 ROC AUC değerleriyle önemli bir başarı elde ettiğini göstermiştir.

Araştırmacılar, ÇOKS algoritmasının eğitim ve sınıflandırma süreçlerinde düşük bir çalışma zamanı maliyetine sahip olması avantajını vurgulamışlardır. Ayrıca, tezde belirtilen ÇOKS algoritmasının mevcut güvenlik tedbirlerine rağmen gelişen dolandırıcı tekniklere karşı etkili bir çözüm sunduğu ifade edilmiştir.

Gelecekteki çalışmalar için öneriler arasında, daha dengeli işlemler içeren veri kümelerinin oluşturulması, çok katmanlı yapay sinir ağı ve derin öğrenme yöntemlerinin kullanılması, ve eğitim ve test sürelerinin optimize edilmesi öne çıkmaktadır. Ayrıca, dolandırıcıların geliştirdiği yeni yöntemlere karşı sürekli olarak adapte olabilecek tekniklerin geliştirilmesi gerektiği belirtilmiştir.[3]

(Yilmaz A., SELİMOĞLU M, şubat 2022) Elde edilen sonuçlara göre, yapılan çalışmada kullanılan Yapay Sinir Ağı algoritmasının Naive Bayes algoritmasından daha başarılı olduğu tespit edilmiştir. Yapay Sinir Ağı, %99,9 doğruluk oranıyla oldukça yüksek bir başarı elde etmiştir, aksine Naive Bayes algoritması %98,2 doğruluk oranı ile biraz daha düşük bir performans göstermiştir.[4]

**VERİSETİ**

Bu projede kullanılan Banksim veri seti, İsveç'teki Blekinge Teknoloji Üniversitesi'nde bir doktora programının parçası olarak oluşturulmuştur.

BankSim, İspanyol bankalarından alınan toplu işlem verilerinin bir örneğine dayanan ajan tabanlı bir simülatördür. Simülatör sahte veri üretme aracıdır ve özellikle dolandırıcılık tespit çalışmalarında kullanılmak üzere tasarlanmıştır. Modeli geliştirmek için istatistiksel ve Sosyal Ağ Analizi (SNA) analizleri kullanılmış ve satıcı ile müşteri arasındaki ilişkiyi tanımlamak ve kalibre etmek için çaba gösterilmiştir.

BankSim'in nihai hedefi, normal ödemeleri bilinen dolandırıcılık göstergeleriyle birleştiren ilgili senaryoları modelleyebilen bir araç haline gelmektir. BankSim tarafından oluşturulan veri kümeleri herhangi bir kişisel bilgi içermez ve meşru ve özel müşteri işlemlerini ortaya çıkarmaz. Böylece, bu veriler akademik kurumların ve diğer şirketlerin dolandırıcılık tespit yöntemleri geliştirmesi ve değerlendirmesi için kullanılabilir. Sentetik veriler ayrıca, kendi verilerine erişimi olanların bu verilerle deney yapmasının daha kolay, hızlı ve ucuz olması avantajına sahiptir ve BankSim'in gerçek verilerin ilgili yönlerine yeterince yaklaştığını iddia etmeyi mümkün kılar.

Bu çalışmada, parametreleri kalibre etmek için BankSim 180 adım (yaklaşık altı ay) boyunca çalıştırılmıştır. Her adımda ortalama üç kartın çalındığı ve yetkisiz işlem sayısının günde yaklaşık iki olduğu hırsızlar tanıtılmıştır. 587443'ü normal ödemeler ve 7200'ü hileli işlemler olmak üzere toplam 594643 giriş oluşturulmuştur. Bu rastgele bir simülasyon olduğu için değerler orijinal verilere tam olarak benzememektedir..

**KULLANILAN TEKNİKLER**

Bu proje çalışması, Jupyter Platformunda ve Python programlama dil ile makine öğrenimi Algoritmaları kullanarak gerçekleştirilmiştir. Projenin temel amacı, bir banka ödeme simülatörü olan BankSim üzerinde, dolandırıcılık tespiti araştırmaları için sentetik veri oluşturmak ve analiz etmektir.

Jupyter, etkileşimli ve açık kaynaklı bir notebook ortamı sunarak veri analizi, görselleştirme ve kod geliştirme süreçlerini birleştiren bir platformdur.

Python, geniş kütüphane desteği ile bilimsel hesaplamalardan veri manipülasyonuna kadar birçok alanda kullanılan güçlü bir programlama dilidir. Bu Proje çalışmasında veri analizi, model geliştirme ve sonuçların görselleştirilimştir.

Makine öğrenimi Algoritmaları, BankSim tarafından üretilen sentetik veri setlerinde normal ödemeleri ve enjekte edilmiş dolandırıcılık belirtilerini ayırt etmek amacıyla uygulanmıştır. Bu süreçte, sınıflandırma modelleri ve istatistiksel analizler kullanılarak dolandırıcı işlemlerin tespiti için bir sistem tasarlanmıştır.

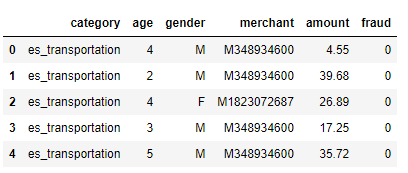
Proje temel bulguları, Jupyter ve Python ile gerçekleştirilen analizlerin, BankSim tarafından üretilen veri setleri üzerinde etkili bir şekilde uygulanabildiğini göstermektedir. Elde edilen sonuçlar, dolandırıcılık tespiti konusunda makine öğrenimi yöntemlerinin kullanılabilirliğini vurgulamakta ve sentetik veri kullanımının avantajlarını ortaya koymaktadır.

**ALGORITMALAR UYGULAMASI**

**1. Adım: Veri Seti Yükleme ve İnceleme**

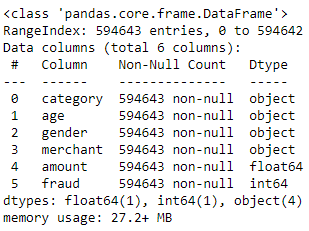
Bu aşamada, Proje çalışılması için temel veri setini yüklenmüştür ve bu veri setini yakından inceleyerek anlamaya çalışılmiştir. Veri setini yüklemek için Python programlama dilinde yaygın olarak kullanılan `pandas` kütüphanesinden faydalanılmiştir. Veri setin "Banksim.csv" adını taşımaktadır.

İlk olarak, veri setinin ilk 5 satırını görüntüleyerek, içerdiği başlıca bilgilere hızlı bir göz atma şansı elde edilmiştir. Bu, veri setindeki değişkenleri ve örnek gözlemleri anlanmasına yardımcı olmaktadır.



Şekil (1) Veri seti ilk 5 satır incelemsi

Ardından, `info()` fonksiyonunu kullanarak veri setinin genel bilgilerini elde edılmiştir. Bu, veri setindeki sütunların türleri, doluluk durumu ve bellek kullanımı gibi önemli bilgileri içerir.



Şekil (2) Veri seti genel bilginin incelemesi

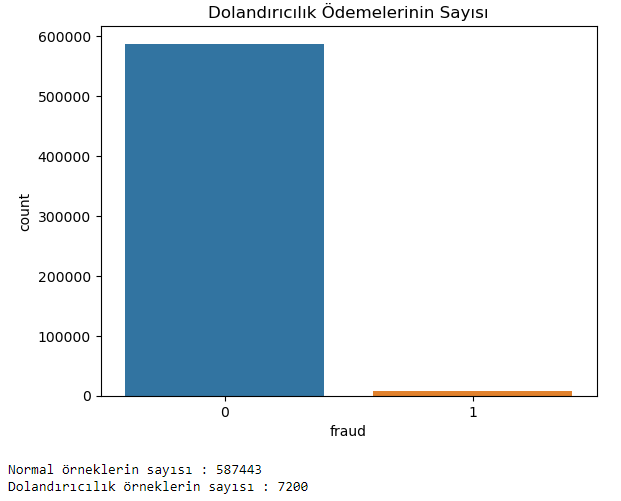
Bu adım, Proje çalışmasının temelini oluşturan veri setini anlama ve önemli özelliklerini belirleme sürecidir. Elde edidiği bu bilgiler, veri seti üzerinde gerçekleştirileceği diğer işlemleri planlanmasını sağlayacaktır.

**2. Adım: Dolandırıcılık ve Dolandırıcılık Olmayan Verilerin Karşılaştırılması**

Proje çalışılmasının üçüncü aşamasında, dolandırıcılık ve dolandırıcılık olmayan verileri içeren iki ayrı veri çerçevesi oluşturulmuştur. Bu adımın temel amacı, bu iki veri kümesi arasındaki farkları inceleyerek dolandırıcılık durumunu etkileyen faktörleri belirlemektir.

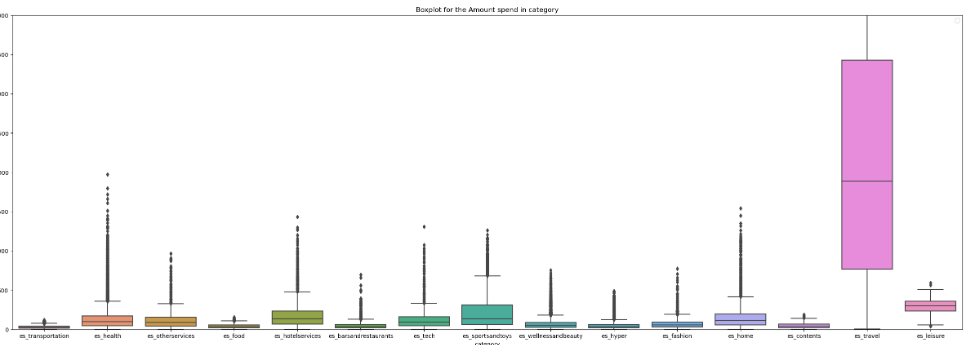
İlk olarak, dolandırıcılık olmayan durumları içeren bir veri çerçevesi (**df\_non\_fraud**) oluşturulmuştur ve aynı şekilde dolandırıcı durumları içeren bir veri çerçevesi (**df\_fraud**) oluşturulmuştur. Bu adım, daha fazla analiz ve görselleştirmeye olanak tanıyarak dolandırıcılık örneklerinin diğer örneklerden ne şekilde farklılaştığını anlanmasına yardımcı olmaktadır.

Daha sonra **Seaborn** kütüphanesini kullanarak dolandırıcılık ve dolandırıcılık olmayan sıparışler arasındaki farkları karşılaştırılmaktadır. Bu noktada iki farklı durumu temsil etmek için seçtiğim iki sütunu karşılaştırarak dolandırıcılık sayısı ile normal satın alma sayısı arasındaki farkı gösteren bir grafik elde edilmiştir.

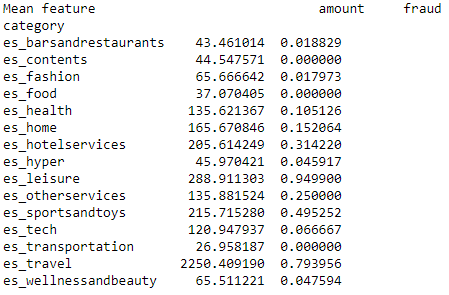


Şekil (3) Dolandırıcılık olayların sayısı Belirtmesi

Dolandırıcılık olaylarının hangi kategoride daha sık gerçekleştiğini incelemek için dolandırıcılık durumlarını içeren veri çerçevesindeki kategorilerin sayısını hesaplayarak Box Plot grafik oluşturulmuştur. Bu grafik, dolandırıcılık olaylarının kategorilere göre dağılımını göstermektedir.



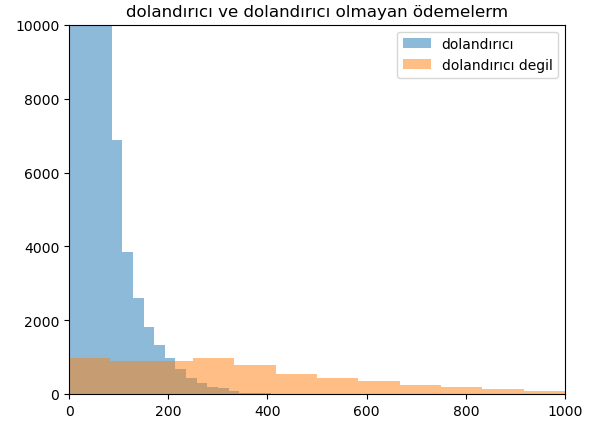
Şekil (4) Dolandırıcılık olaylarının hangi kategoride daha sık gerçekleştiğini incelemesi



Şekil (5) Dolandırıcılık olaylarının hangi kategoride daha sık gerçekleştiğini incelemesi

Yukarıdaki kategoriye göre ortalama tutarı ve dolandırıcılık yüzdesini görulmektedir. sonuçlara göre, eğlence ve seyahat, dolandırıcılar için en çok tercih edilen kategorilerdir. Dolandırıcılar, insanların ortalama olarak daha fazla harcama yaptığı kategorileri seçmiş gibi görünümektedır.

Aşadaki Grafik Dolandırıcılık durumlarının hangi miktarlarda daha sık gerçekleştiğini incelenmktedir

s

Şekil (6) : Dolandırıcılık olaylarının hangi miktarlarda daha sık gerçekleştiğini incelemesi

Genellikle daha az miktarda işlem yapmış gibi görünmektedir. Daha düşük miktarlı işlemler, dolandırıcıların dikkat çekmeden ve sistemi kötüye kullanmadan daha kolay olabilir. Bu nedenle, dolandırıcılar genellikle küçük miktarlı işlemleri tercih edebilidiği görülmüştür.

**3. Adım: Veri Dönüştürme ve Bağımsız/Bağımlı Değişkenlerin Belirlenmesi**

Bu adımda, veri dönüştürme sürecini kolaylaştırmak için nesne türündeki sütunları kategorik hale getirilmiştir. Bu, makine öğrenimi modellerinin kullanımı için veri setini uygun formata getirme sürecidir. Daha sonra, bağımsız değişkeni (X) ve bağımlı/hedef değişkeni (y) tanımlanmaktadır.

**4. Adım: Örnekler Dengeli Hale Getirilmesi**

Daha önce görüldüğü gibi, dolandırıcılık ve dolandırıcılık olmayan olaylar arasında oldukça bir sayı farkı bulunmaktadır.

Dolandırıcılık verileri dengesiz olacaktır, bu durumu hem grafikte gördüğümüz gibi hem de örnek sayısından anlaşılıbılır. Veri setini dengelemek için oversample veya undersample tekniklerini uygulamak mümkündür. Oversampling, azınlık sınıfın örnek sayısını artırarak azınlık sınıfından örnekler oluşturmayı içerir. Undersampling ise çoğunluk sınıfındaki örnek sayısını azaltarak rastgele noktaları seçmeyi içerir, böylece azınlık sınıfı ile eşit hale gelir. Her iki işlem de bazı risklere sahiptir: Oversample, kopya veya benzer veri noktaları oluşturacaktır, bu da dolandırıcılık tespiti durumunda bazen yardımcı olmayabilir çünkü dolandırıcı işlemler farklılık gösterebilir. Undersample ise veri noktalarını ve dolayısıyla bilgiyi kaybetmemize neden olur.

Proje kapsamında veriyi dengelemek için SMOTE (Synthetic Minority Over-sampling Technique) adlı bir oversampling tekniği uygulanmiştir. SMOTE, azınlık sınıfından yeni veri noktaları oluşturmak için komşu örnekleri kullanır, bu nedenle oluşturulan örnekler tam kopyalar olmasa da benzerdir.

**5. Adım: Makine Öğrenimi Algoritmalarının Uygulanması**

Dengelediğimiz veri seti üzerinde çeşitli makine öğrenimi algoritmalarını uygulanılmiştir. Bu algoritmalar arasında İleri Beslemeli Sinir Ağı (Artificial Neural Network), Gaussian Naive Bayes, K-Nearest Neighbors (KNN) veXGBoost bulunmaktadır. Bu algoritmalar, dolandırıcılık tespiti için kullanılan farklı prensiplere dayanan modellerdir. Uygulanan algoritmaların performansını değerlendirmek için eğitim ve test veri setlerini kullanarak doğruluk (accuracy), confusion matrix, ve sınıflandırma raporu gibi metrikleri incelenmiştir.

**SONUÇ**

Aşadaki tablo kullanılan Algoritmalar Görülmektedir

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Accuracy** | **Percision** | **Recal** | **F1 Score** |
| **Ann** | **0.974** | **0.972** | **-** | **-** |
| **Gnb** | **0.92** | **0.92** | **0.92** | **0.92** |
| **Knn** | **0.99** | **0.99** | **0.99** | **0.99** |
| **XGBoost** | **0.99** | **0.99** | **0.99** | **0.99** |

Tablo (1) : Kullanılan Algoritmalar Sonuçları

İleri Beslemeli Sinir Ağı (ANN) modeli, eğitim sırasında kaybın azalması ve doğruluğun artmasıyla iyi bir performans sergilemiştir. Model, eğitimde %97.25 doğruluk elde etmiş ve doğrulama setinde %97.42 doğruluk sağlamıştır.

Gaussian Naive Bayes modeli ise, 0 sınıfı için 0.92, 1 sınıfı için 0.92 olan bir F1 skoruna sahiptir. Hassasiyet ve duyarlılık değerleri sırasıyla 0.91, 0.93 ve 0.93, 0.91'dir. Confusion matrix, modelin 1 sınıfındaki bazı örnekleri kaçırma eğiliminde olduğunu göstermektedir.

K-Nearest Neighbors modeli ise, yüksek doğruluk oranı (%99) ve düşük hata oranları ile oldukça başarılı bir performans sergilemiştir. Hassasiyet ve duyarlılık değerleri yüksek olup, 1 sınıfındaki örnekleri kaçırma oranı düşüktür.

Bu sonuçlara dayanarak, K-Nearest Neighbors modelinin diğer iki modele göre daha iyi bir performans gösterdiği söylenebilir.

XGBoost (eXtreme Gradient Boosting) modeli ise, 0 sınıfı için 0.99 hassasiyet ve 0.98 duyarlılık, 1 sınıfı için ise 0.98 hassasiyet ve 0.99 duyarlılık elde ederek yüksek performans sergilemiştir. F1-skorları sırasıyla 0.99, doğruluk oranı ise %99 olarak belirlenmiştir. Bu sonuçlar, XGBoost'un dolandırıcılık tespiti konusunda etkili bir model olduğunu göstermektedir.

**KAYNAKÇA**

[1] ÖZBAY E. , “Finans Sektöründe Veri Madenciliği ile Dolandırıcılık Tespiti” (Doctoral Dissertation, Selçuk Üniversitesi Fen Bilimleri Enstitüsü), 2007

[2] ÖZMEN E. P. , ÖZCAN T. , “Finans Sektöründe Dolandırcılık Tespiti Üzerine Melez Sınıflandırma ve Regresyon Ağacı Uygulaması”

[3] AYDIN T. , Dal D. ,” Yapay Zeka Destekli ÇOKS Yöntemi ile Kredi Kartı Sahtekarlığının Tespiti” ,Mayıs 2021

[4] Yilmaz A., SELİMOĞLU M , “KREDİ KARTI DOLANDIRICILIK TESPİTİNİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE TAHMİN EDİLMESİ”,şubat 2022